UNIVERSITY
OF TRENTO

DIPARTIMENTO DI INGEGNERIA E SCIENZA DELL'INFORMAZIONE

38123 Povo – Trento (Italy), Via Sommarive 14
http://www.disi.unitn.it

GENETICALLY-DESIGNED ARBITRARY LENGTH ALMOST DIFFERENCE SETS

G. Oliveri, M. Donelli, and A. Massa

January 2011

Technical Report # DISI-11-078

Genetically-Designed Arbitrary Length Almost Difference Sets

G. Oliveri, M. Donelli, and A. Massa

Almost Difference Sets (ADSs) have important applications in cryptography, coding theory, and antenna array thinning. In this letter, a new approach is proposed to derive ADSs of arbitrary lengths. Such a technique recasts the ADS design as a combinatorial optimization problem successively solved by means of a suitable binary Genetic Algorithm. New ADSs are derived to assess the effectiveness of the proposed approach.

Introduction: In recent years, the design of binary sequences with three-level autocorrelation [e.g., the so-called Almost Difference Sets (ADSs)] has gained considerable attention because of several applications in cryptography [1], coding [2]-[4], and antenna array synthesis [5]. Several construction techniques have been already developed [6]-[9], and even large repositories are now available [10]. However, the fact that ADS sequences of arbitrary length are (at present) not available [6]-[10] is a limitation for their use in real-world problems. As a matter of fact, since ADS synthesis techniques are usually based on the *cyclotomy* property [9], they generate sequences characterized by specific cyclotomic numbers and not with arbitrary length.

In this letter, a new method is proposed for the synthesis of ADS sequences of arbitrary length. The approach reformulates the ADS design in terms of a combinatorial optimization problem where the cost function quantifies the misfit between the autocorrelation of a binary sequence and the three-valued function of the ADSs. A binary genetic algorithm (GA) [11][12] is used to minimize such a cost function because of its “hill climbing” features and its ability to sample in a

very effective fashion the binary solution space [11]. Representative numerical results are analyzed to give some indications on the effectiveness of the proposed approach and explicit expressions of new ADS sequences of different lengths are derived.

ADS Synthesis as a Combinatorial Optimization Problem: Let us consider a binary sequence $\underline{b} = \{b(n); n = 0,...,N-1\}$ of length N whose cyclic autocorrelation function is defined as

$$A(\tau) = \sum_{n=1}^{N} b(n) b[(n+\tau) \bmod N] \tag{1}$$

where $\tau = 0,...,N-1$ and $b(n) \in \{0,1\}$.

A (N,K,Λ,t)-ADS is a binary sequence characterized by an autocorrelation function [6]-[9] given by

$$A_{ADS}(\tau) = \begin{cases} K & \tau = 0 \\ \Lambda & for\ t\ values\ of\ \tau \\ \Lambda + 1 & otherwise \end{cases}. \tag{2}$$

In order to determine an ADS of arbitrary length N, a suitable optimization problem is formulated and then solved by a means of a GA-based iterative procedure [11][12]. Starting from an initial ($i = 1$) population composed by a set of P randomly-generated binary sequences, $\underline{b}_p^i$, $p = 0,...,P-1$, the trial solutions iteratively (i being the iteration index) undergo the genetic evolution through selection, crossover (with probability χ), and mutation (with probability μ) until a suitable solution $\underline{b}_{opt}$ is found. The degree of optimality of each trial sequence is evaluated by means of the following fitness function

$$F(\underline{b}_p^i) = \alpha(L_p^i - 3) + \beta R_p^i \tag{3}$$

where L_p^i is the number of levels of the autocorrelation function of $\underline{b}_p^i$ and R_p^i is the number of τ values for which $A_p^i(\tau)$ differs from (2). Moreover, α and β are user-defined real weights. The iterations stop when either an ADS sequence is found [i.e., $F(\underline{b}_{opt}) = 0$] or the maximum number of iterations I is reached ($i = I$).

It is worth to notice that, since (3) measures the "similarity" between $A_p^i(\tau)$ and the autocorrelation function of the ADSs, the optimization technique synthesizes the optimal binary sequences by iteratively approaching the desired three-level autocorrelation, thus avoiding the constraints on the sequence length of state-of-the-art generation methods [6]-[9].

Numerical validation: For a preliminary assessment, the following parameter setup has been chosen throughout the numerical validation: $\chi = 0.9$, $\mu = 0.01$, $\alpha = 10^{-2}$, $\beta = 10^{-4}$, and $P = N$.

The first example deals with the synthesis of an ADS of length $N = 24$. Although the dimension of the solution space is not negligible ($U = 2^{24} \cong 1.7 \times 10^7$, U being the number of binary sequences), very few iterations ($i_{opt} = 8$) are enough to reach a binary sequence with a three-level autocorrelation [Fig. 1(*a*)], which (to the best of authors' knowledge) corresponds to a new ADS, namely the $(24,8,2,13)$-ADS in Fig. 1(*b*).

A more complex synthesis problem has been addressed in the successive example concerned with the design of an ADS with $N = 90$. The $(90,8,0,59)$-ADS in Fig. 2(*a*), whose autocorrelation is given in Fig. 2(*b*), has been obtained after $i_{opt} = 108$ iterations as indicated by Fig. 2(*c*) where the iterative evolution of the optimal value of the fitness function, $\Omega^i = \min_p \left\{ F(\underline{b}_p^i) \right\}$, is shown. For

completeness, let us consider that the whole GA-based minimization required the evaluation of about 9000 trial solutions to sample a search space of $U = 1.23 \times 10^{27}$ sequences.

The last example is concerned with a higher-dimension problem, being $N = 200$. As expected, the number of iterations to identify $\underline{b}_{opt}$ increases with respect to previous experiments [Fig. 3(*a*)] even though less than $i_{opt} = 350$ are needed to fit the three-level autocorrelation function in Fig. 3(*b*). The $(200,8,0,143)$-ADS sequence is shown in Fig. 3(*c*).

Conclusions: In this letter, a GA-based technique has been proposed as a new methodological tool for designing ADS sequences of arbitrary length. The original synthesis has been reformulated as a combinatorial optimization. Towards this end, a suitable fitness function exploiting the autocorrelation properties of ADSs has been introduced and minimized by means of a GA-based iterative procedure. Selected numerical results have been reported to give a preliminary assessment of the capabilities and effectiveness of the proposed approach.

References:

1 CARLET C., and DING C., “Highly nonlinear mappings,” J. Complexity, vol. 20, no. 2, pp. 205-244, Apr. 2004.

2 DING C., and FENG T., “Codebooks from almost difference sets,” Des. Codes Cryptography, vol. 46, no. 1, pp. 113-126, 2008.

3 DING C., “Complex codebooks from combinatorial designs,” IEEE Trans. Inf. Theory, vol. 52, no. 9, pp. 4229-4235, Sep. 2006.

4 DING C., and FENG T., “A generic construction of complex codebooks meeting the welch bound,” IEEE Trans. Inf. Theory, vol. 53, no. 11, pp. 4245-4250, Nov. 2007.

5 OLIVERI G., DONELLI M., and MASSA A., "Linear array thinning exploiting almost difference sets," IEEE Trans. Antennas Propag., in press.

6 DING C., HELLESETH T., and LAM K. Y., “Several classes of binary sequences with three-level autocorrelation,” IEEE Trans. Inf. Theory, vol. 45, no. 7, pp. 2606-2612, Nov. 1999.

7 DING C., HELLESETH T., and MARTINSEN H., "New Families of Binary Sequences with Optimal Three-Level Autocorrelation," IEEE Trans. Inf. Theory, vol. 47, no. 1, pp. 428-433, Jan. 2001.

8 ARASU K. T., DING C., HELLESETH T., KUMAR P. V., and MARTINSEN H. M., "Almost difference sets and their sequences with optimal autocorrelation," IEEE Trans. Inf. Theory, vol. 47, no. 7, pp. 2934-2943, Nov. 2001.

9 ZHANG Y., LEI J. G., and ZHANG S. P., "A new family of almost difference sets and some necessary conditions," IEEE Trans. Inf. Theory, vol. 52, no. 5, pp. 2052-2061, May 2006.

10 ELEDIA Almost Difference Set Repository (*http://www.eledia.ing.unitn.it*)

11 HAUPT, R., and WERNER, D. H.: *Genetic algorithms in electromagnetics*. Hoboken, NJ: Wiley, 2007.

12 CAORSI S., LOMMI A., MASSA A., PASTORINO M., “Peak sidelobe level reduction with a hybrid approach based on GAs and difference sets,” IEEE Trans. Antennas Propag., vol. 52, no. 4, pp. 1116-1121, Apr. 2004.

Authors' affiliations:
G. Oliveri, M. Donelli, and A. Massa
ELEDIA Research Group
Department of Information Engineering and Computer Science,
University of Trento, Via Sommarive 14, 38050 Trento – ITALY

Corresponding Author:
A. Massa
ELEDIA Research Group
Department of Information Engineering and Computer Science
University of Trento
Via Sommarive 14
38050 Trento
ITALY
E-mail: andrea.massa@ing.unitn.it

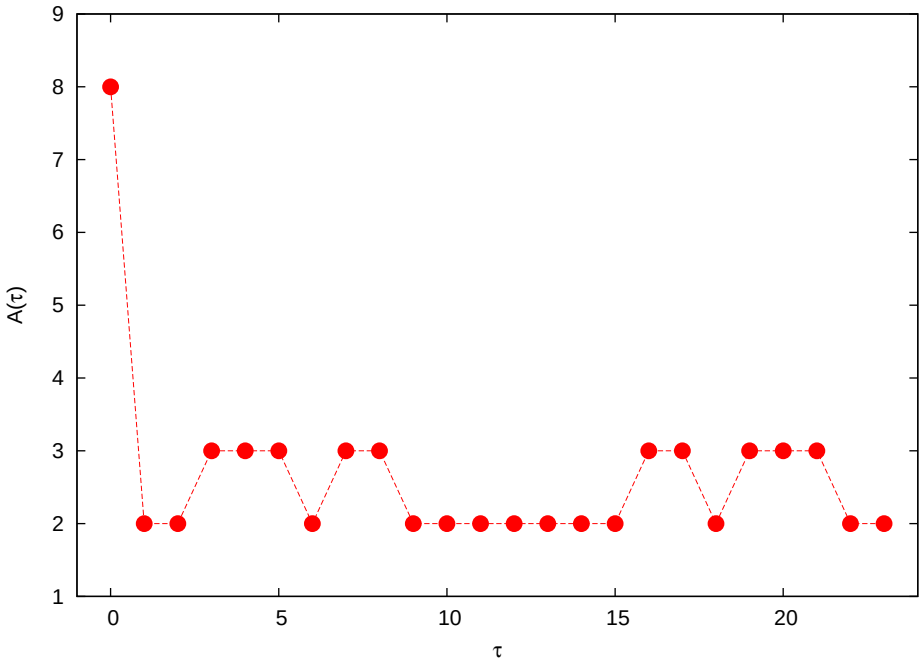
Figure captions:

Fig. 1 - *Numerical validation* ($N = 24$) - (24,8,2,13)-ADS: (*a*) the autocorrelation function, $A_{opt}(\tau)$, and (*b*) the binary sequence, $\underline{b}_{opt}$.

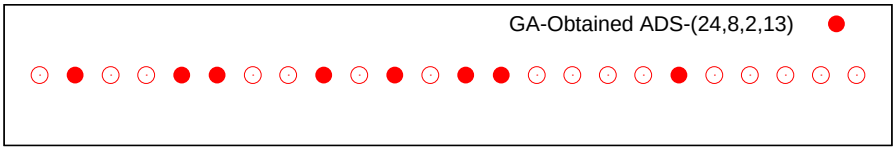
Fig. 2 - *Numerical validation* ($N = 90$) – (90,8,0,59)-ADS: (*a*) the binary sequence, $\underline{b}_{opt}$, (*b*) the autocorrelation function, $A_{opt}(\tau)$, and (*c*) the optimal fitness, Ω^i, versus the iteration index, *i*.

Fig. 3 - *Numerical validation* ($N = 200$) - (200,8,0,143)-ADS: (*a*) the optimal fitness, Ω^i, versus the iteration index, *i*, (*b*) the autocorrelation function, $A_{opt}(\tau)$, and (*c*) the binary sequence, $\underline{b}_{opt}$.

Figure 1

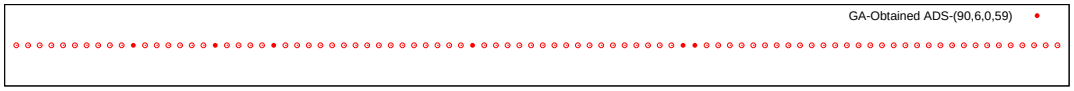


(*a*)

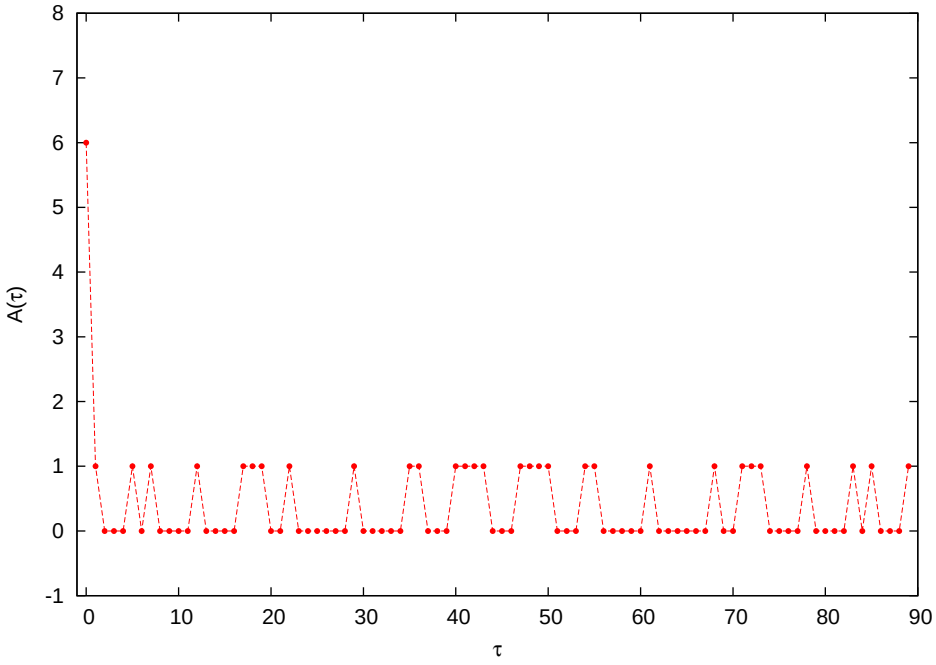


(*b*)

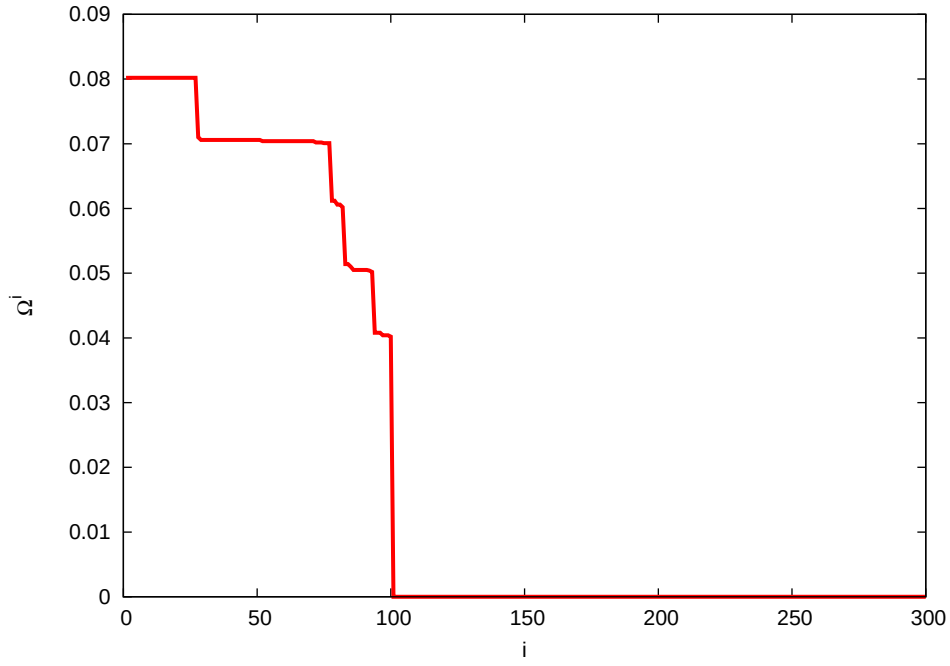
Figure 2


GA-Obtained ADS-(90,6,0,59)

(*a*)

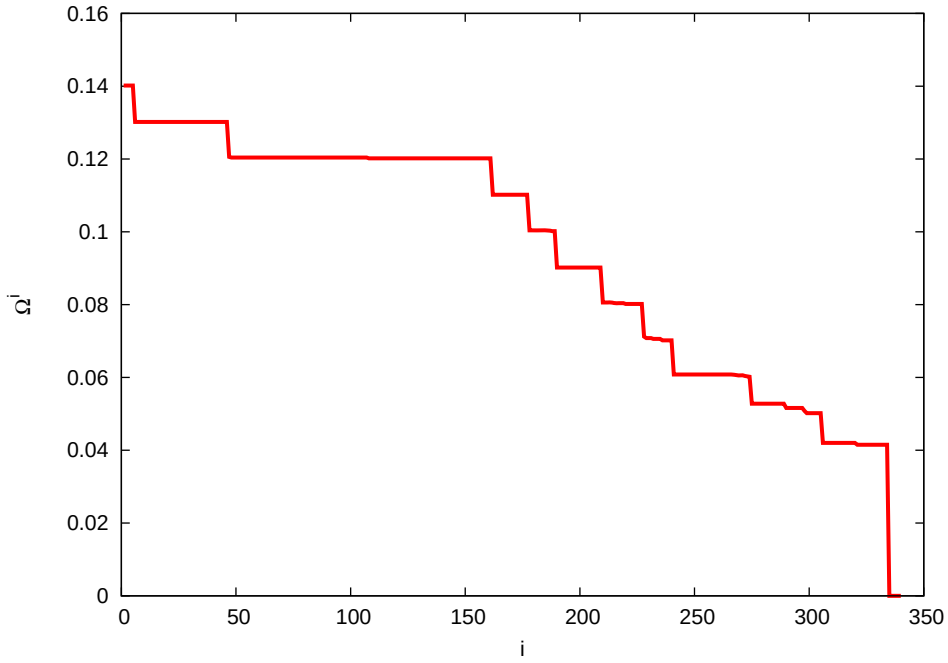

A(τ)
τ
8
7
6
5
4
3
2
1
0
-1
0
10
20
30
40
50
60
70
80
90

(*b*)

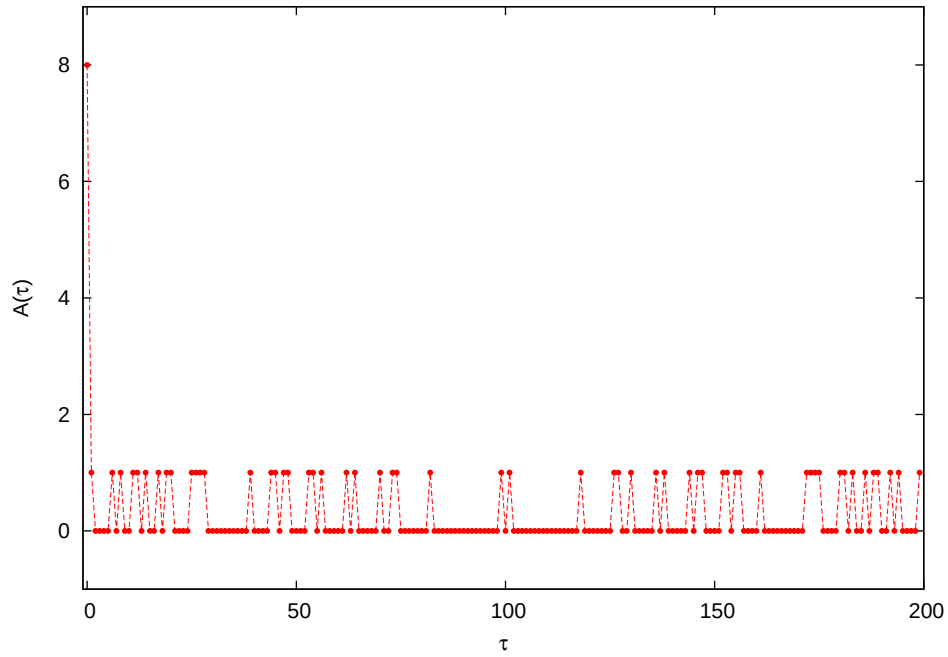

Ω^i
i
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
50
100
150
200
250
300

(*c*)

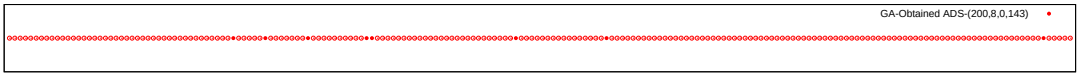
Figure 3


0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Ω^i
0
50
100
150
200
250
300
350
i

(*a*)


8
6
4
2
0
A(τ)
0
50
100
150
200
τ

(*b*)


GA-Obtained ADS-(200,8,0,143)

(*c*)